Construing Science in the Quest for "*Ipse Dixit*": A Comment on Sanders and Cohen

Joe S. Cecil[*]

But what's a judge to do?

Professor Sanders offers a convincing argument that some degree of paternalism is appropriate in guarding against misinterpretation of complex scientific and technical evidence by lay jurors, especially when jurors must consider statistical and probabilistic evidence. His review of research suggest that jurors often respond to complex evidence by engaging "in peripheral processing when accessing expert testimony and that peripheral cues [e.g., credentials and demeanor] take on added significance as the scientific issues in the case become more complex."[1] The fact that such evidence is embedded in an adversarial presentation intended to favor the presenting side makes it even less accessible. Sanders concludes that if the purpose of the trial is to ascertain the truth, some restriction on admissibility of evidence is appropriate to maximize the likelihood of arriving at a correct answer.

Professor Cohen complicates the issue considerably by pointing out that the legal system, in making a gatekeeping determination, must accommodate values that go beyond the values of the sciences. More specifically, the legal system weighs the cost of errors differently and "the fact that a 'mainstream' scientist would not testify as to a particular conclusion does not necessarily mean that the same conclusion is valueless or 'junk science' for the purposes of law."[2] Cohen is particularly concerned that in focusing on scientific "knowledge" the courts will misunderstand "how data (i.e., evidence) is utilized to develop statements of 'knowledge' in scientific

---

[*] Senior Research Associate, Federal Judicial Center. The views expressed in this article are those of the author and do not necessarily represent the views of the Federal Judicial Center.

[1] Joseph Sanders, *The Paternalistic Justification for Restrictions on the Admissibility of Expert Evidence*, 33 SETON HALL L. REV. 881, 913 (2003).

[2] Neil B. Cohen, *The Gatekeeping Role in Civil Litigation and the Abdication of Legal Values in Favor of Scientific Values*, 33 SETON HALL L. REV 943, 945 (2003).

communities consistent with the norms of those communities—norms that differ fundamentally from the norms of civil litigation."[3]

One focal point of the tension between law and science arises in the many circumstances in which science would conclude that a relationship is "'suggested but not proven.'"[4] This typically occurs when a research study finds an effect that points in the direction of an association, but cannot rule out the possibility that the observations are a product of chance, perhaps due to a small sample size. In such a circumstance, the researchers will suspend judgment, an option that is not available to legal factfinders. Cohen points out that the legal system will often translate such uncertainty into a decision to exclude the evidence, to the detriment of the plaintiff who has the burden of proof. In doing so, the courts may inadvertently give greater weight to errors favoring defendants' interests rather than errors favoring plaintiffs' interests and thereby distort the values of the civil justice system, which assumes that such errors should be given equal weight.[5] Using epidemiology evidence as an example, Cohen demonstrates that the conservative values implicit in declaring the existence of an effect place an awesome barrier in the path of the plaintiff who wishes to present expert evidence to the jury.

Taken together, the positions of Professors Sanders and Cohen present a tall order for judges. In making a gatekeeping decision they must not only anticipate the shortcomings of jurors in assessing complex evidence and exercise their authority in a way that strengthens the accuracy of the process, but they also must be alert to the tension in the values of science and law and construe "scientific knowledge" in a way that does not permit the implicit values of science to override the implicit values of civil litigation.

If only it were so easy. The problem is that there is not just one science and not one scientific method.[6] Perhaps the values of the civil litigation system can be characterized in a uniform manner across a multitude of cases, but the values of science vary across the individual disciplines. Each one has its own norms and standards that vary greatly in the rigor they impose in declaring a finding to be "scientific knowledge."

---

[3] *Id.* at 946.

[4] *Id.* at 950.

[5] *Id.* at Part II.

[6] Sheila Jasanoff has referred to this as the "myth of the scientific method." Sheila Jasanoff, Hidden Experts: Judging Science after *Daubert* 9-12 (unpublished manuscript, on file with author).

In a typical gatekeeping decision a judge is likely to find practitioners of many sciences, each offering information that conforms to the standards of his or her narrow scientific specialty. The differences among scientific disciplines are evident in the pedigree of experts as they line up at trial. Plaintiffs often favor expert toxicologists, who typically rely on animal studies for estimating the effects of exposure on living organisms. Although toxicologists may face difficult problems of extrapolation across species and dosage levels, they are comforted by the greater control over external exposures permitted by laboratory studies.[7]

Defendants often favor epidemiologists, whose broad-scale studies of the effects of various exposures on humans avoid the problems of extrapolation across species and dosage levels. Epidemiology studies, however, are difficult to execute and subject to external influences that are uncontrolled by the research design, influences that may compromise the confidence with which the findings of the study may be attributed to a specific exposure.[8] As Cohen points out, epidemiology studies are conservative in nature, willing to declare a causal relationship only when the statistical tests speak conclusively in ruling out chance as an alternative explanation, and when a series of additional restrictive conditions are met.[9]

By focusing on epidemiology, Professors Sanders and Cohen consider only one scientific value system. Other sciences endorse other value systems, resulting conflicts among scientists from different disciplines requires judges to resolve issues that the sciences themselves have left unresolved. When courts are faced with conflicting sciences, they must choose among them, taking cognizance of some forms of scientific knowledge and dismissing others. Such a choice is unavoidable. In doing so, courts often invoke the language of science, thereby obscuring the fact that they are imposing legal values and standards. Consequently, the language of science in the courtroom can misrepresent both the science and the law.

---

[7] Bernard D. Goldstein & Mary Sue Henifin, *Reference Guide on Toxicology*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 405 (2d ed. 2000).

[8] David P. Rall et al., *Alternatives to Using Human Experience in Assessing Health Risks*, 8 ANN. REV. PUB. HEALTH 355, 362-63 (1987).

[9] Cohen, *supra* note 2, at 951-52 ("Most commonly, the epidemiologist will decline to characterize the data as showing a particular relationship between A and B unless the probability that the relationship could have occurred by chance even in the absence of the observed relationship is quite low. The most common threshold for such a probability is 5%."). See also criteria for assessing whether an association reflects a causal relationship in Michael D. Green et al., *Reference Guide on Epidemiology*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 376-77 (2d ed. 2000).

The Supreme Court has not offered much guidance to lower courts charged with sorting out conflicting areas of science. Repeatedly the Supreme Court has spoken as though there is a single standard of scientific values that lead to a uniform standard of "scientific knowledge." In *Daubert*, Justice Blackmun urged judges to measure the admissibility of expert testimony against the extent to which the testimony was based on "scientific validity,"[10] and offered four non-exclusive factors to consider in making this assessment. While the Court expressly limited its holding to scientific evidence,[11] it chose not to delineate how these factors would apply across the range of sciences that are proffered in court.

The difficulty of establishing a uniform standard of evidential reliability across competing sciences became apparent in the Court's next declaration on scientific evidence. In *General Electric v. Joiner*,[12] the Supreme Court indicated that:

> [n]othing in either Daubert or the Federal Rules of Evidence requires a district court to admit opinion evidence which is connected to the existing data only by the *ipse dixit* of the expert. A court may conclude that there is simply too great an analytical gap between the data and the opinion proffered. This is what the District Court did here and we hold that it did not abuse its discretion in so doing.

But what constitutes *too great* an analytical gap? Sanders and his colleagues nailed the question (but not the answer, in my view) when they asked "[h]ow good is good enough?"[13] In setting a threshold for admissibility, the law must endorse a single answer, while the sciences tolerate a multitude of answers.[14]

The dilemma of competing sciences was sharpened in *Kumho Tire Co. v. Carmichael*,[15] which extended the gatekeeping role beyond the sciences to all expert testimony. Although the *Daubert* factors always may be considered, Justice Breyer noted that other factors may provide a more suitable standard for assessing such testimony. The Supreme Court indicated that all expert witnesses should employ "in

---

[10] Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 590 n.9 (1993).

[11] *Id.* at n.8.

[12] 522 U.S. 136, 146 (1997)

[13] David L. Faigman et al., *How Good is Good Enough?: Expert Evidence Under* Daubert *and* Kumho, 50 CASE W. RES. L. REV. 645 (2000); Sanders, *supra* note 1, at 939.

[14] Any notion of a monolithic science can be dispelled by examining the 271 scientific societies affiliated the American Association for the Advancement of Science. *See* http://www.aaas.org/about/affiliates.shtml (last visited Mar. 24, 2003).

[15] 526 U.S. 137 (1999).

the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field."[16] In effect, this decision tethered the standard for admissibility of testimony to standards of professional practice, and implicitly recognized that this standard of validity may vary within the sciences.

Taken together, these cases leave judges struggling to resolve conflicting evidence from different areas of science, all of which employ "the same level of intellectual rigor" that is appropriate to the individual area of science, but which tolerate different degrees of extrapolation in transcending the "analytical gap" between scientific findings and the circumstances of the individual litigation.

I. THE REVEALING CASE OF *SOLDO V. SANDOZ*

An especially revealing instance of the dilemma posed by differing standards among sciences is found in the recent district court decision in *Soldo v. Sandoz*.[17] Like the *Kuhn* case cited by Sanders, *Soldo* is another instance of personal tragedy and uncertain scientific attribution. A young mother sustained an intracranial hemorrhage and resulting stroke soon after giving birth. She claimed that the stroke was a consequence of her ingestion of Parlodel, a drug manufactured and marketed by the defendant to prevent lactation. Since a heightened risk of stroke occurs in the postpartum phase of pregnancy,[18] among the tasks facing the court was to determine if the risk of stroke among women who had taken Parlodel following pregnancy exceeded the heightened base rate of stroke among women in the postpartum phase of pregnancy. No meaningful epidemiology studies were available because the incidence of stroke was rare and the FDA removed the product from the market in 1995 following reports of adverse health effects in post-partum women.

The defendant pharmaceutical manufacturer challenged the testimony of the plaintiff's two primary experts, Drs. Kulig and Petro, who were prepared to testify that they had excluded alternative possible causes of the stoke and were confident to a reasonable degree of medical certainty that ingestion of Parlodel caused the stroke.

Soldo merits close examination because the court also appointed three distinguished scientists from three different disciplines to

---

[16] *Id.* at 152.

[17] Soldo v. Sandoz Pharm. Corp., 244 F. Supp. 2d 434, Civ. No. 98-1712, 2003 WL 355931 (W.D. Pa. Jan. 13, 2003).

[18] *Id.* at *10.

advise the court as independent experts concerning the scientific acceptability of the plaintiff's expert evidence. The three court appointed experts were Dr. David A. Savitz, Professor and Chair of the Department of Epidemiology at the University of North Carolina; Dr. William J. Powers, Professor in the Departments of Neurology and Radiology at the Washington University School of Medicine; and Dr. David Flockhart, Chief of the Division of Clinical Pharmacology at the Indiana University School of Medicine.[19] The use of such distinguished court-appointed experts allows a clear assessment of the manner in which different areas of science assess similar evidence, unfiltered by the adversarial interests of the parties. Unlike other recent instances in which courts employed panels of appointed experts,[20] these court-appointed experts worked independently, remaining unaware of the identity or areas of specialization of the other appointed experts. This independence permits a clear assessment of the perspective of the three areas of science without the dilution that might arise if the experts were expected to collaborate on a single report.

In a lengthy order, the court sought the advice of the appointed experts on "whether the methodology or technique employed by plaintiff's medical witnesses, Dr. Kenneth Kulig and Dr. Dennis Petro, in formulating their opinions [was] scientifically reliable and whether the methodology or technique [could] be properly applied to the facts of this case."[21] The court recited an expanded version of the *Daubert* factors, noting that published studies are not required to establish reliability; that "differential diagnosis and temporal analysis, properly performed, would generally meet the factors"; and that there must be a proper "fit" between the expert's opinion and the facts of the case.[22]

---

[19] The experts were selected with the assistance of the Registry of Independent Scientific and Technical Advisors, a program of the Private Adjudication Center at Duke University School of Law. *See* http://www.law.duke.edu/pac/registry/index.html. The executive director of the program nominated candidates to serve as court appointed experts after consulting with the court regarding the areas of expertise that were required. The executive director also served as a channel of communication between the court and the experts; no direct communication with the court was permitted.

[20] Laural L. Hooper et al., *Assessing Causation in Breast Implant Litigation: The Role of Science Panels*, 64 LAW & CONTEMP. PROBS. 139 (2001).

[21] Soldo vs. Sandoz Pharm. Corp., Civ. No. 98-1712, Order of the Court, Exhibit I, Court's Instructions to Experts Appointed Pursuant to Federal Rule of Evidence 706, 1 (W.D. Pa. Mar. 26, 2001).

[22] *Id.* at 2-3. The court continued, "There is no 'fit' where there is simply too great an analytical gap between the data and the opinion offered, as when an expert offers animal studies showing one type of cancer in laboratory mice to support

More specifically, the court asked each expert to respond to the following questions with regard to Drs. Kulig and Petro:

> 1. With a reasonable degree of medical certainty, is the methodology or technique employed by [the plaintiff's expert] in opining that Parlodel can cause stroke and that Parlodel caused plaintiff's intracerebral hemorrhage scientifically reliable?
>
> 2. Can the methodology or technique employed by [the plaintiff's expert] be applied to the facts at issue?
>
> 3. If your answers to [either of these questions] are in the negative, to what extent, if any, should any of your opinions be considered subject to sufficient genuine dispute as would permit other persons, generally qualified in your field of expertise, to express opinions that, though contrary to yours, would likely be viewed by others in the field as representing legitimate and reasonable disagreement within your profession?[23]

The court provided each expert with transcripts of a seven-day *Daubert* hearing and related exhibits, and instructed them not to communicate with the parties or each other.

A. The Expert Reports

The reports filed by the three court-appointed experts demonstrate the diverse values that differing scientific disciplines employ in assessing an identical body of evidence.[24] All acknowledged that the plaintiff's experts had made the most of the data that existed, but differed in the extent to which they regarded such testimony as meeting the standards of "scientific knowledge."

Dr. David Savitz offered the most demanding standard for finding information to be sufficiently reliable to be deemed scientific knowledge:

> because the information is so indirectly applicable and hypothetical in nature, the application of it to form an opinion is not a "scientifically reliable" process. The linkage between those shreds of potentially relevant information and the opinion that results is so murky that it is very difficult to see how the evidence leads to the opinions that are offered. Applying any reasonable standards of scientific evidence as the basis for drawing a

---

causation of another type in humans." *Id.* at 3.

[23] *Id.* at 5-6. This third question follows the format of a question to court-appointed experts in the breast implant multidistrict litigation. *See* Hooper et al., *supra* note 20, at 158.

[24] Since the instructions to the court-appointed experts and the reports of those experts remain unpublished, I have provided extensive citation to the language of the instructions and reports.

> conclusion leads to the judgment that we do not know enough to offer an opinion on this matter that is reasonably well grounded in science.[25]

Dr. Savitz, an epidemiologist, was especially critical of the absence of human studies and unwilling to extrapolate findings across species, declaring that "[s]ome form of epidemiologic or clinical evidence, even if flawed and incomplete, is needed for drawing inferences about general causation in making a judgment about Parlodel and intracerebral hemorrhage."[26] Dr. Savitz acknowledged that causal attributions may be made without such studies if the causal pathway is clear; he then continued, "[i]n the absence of clinical or epidemiologic research, it would require a tremendous amount of indirect evidence to reach the point that even in the absence of research, the linkage is 'obvious' in the way that the tornado leading to injury is obvious."[27]

With respect to the third question involving the possibility of legitimate and reasonable disagreement within the profession regarding such views, Dr. Savitz acknowledged that views vary across sciences:

> [The] vast majority of scientists who routinely consider these sorts of evidence (epidemiologists, researchers in clinical medicine) would agree with my general conclusions. If forced to guess the proportion, I would estimate 80% of my peers would concur. Those who study basic mechanisms of disease causation (physiologists, pharmacologists, toxicologists) might well dispute my views in that the plausibility based on those lines of evidence is more supportive of the potential for causality. The counterargument to my view is that the diverse threads based on mechanism of actions for Parlodel, analogy to other agents in the same broad category of drug, and temporal linkage of the medication and the illness can be integrated scientifically into a scientifically reliable conclusion. However, essentially all scientists recognize that when the issue is the causation of clinical disease in humans, there is a sizable gap between what is plausible based on indirect evidence and what is proven based on clinical and epidemiologic studies. In the chain of reasoning, most scientists would likely share the view that leaping across the huge gulf of critical data moved the person making the inference beyond the

---

[25] Memorandum from David A. Savitz, to Judge Donald J. Lee 4 (Sept. 31, 2001) (on file with author) [hereinafter Savitz].

[26] *Id.* at 3.

[27] *Id.* at 4.

scope of scientific assessment.[28]

Dr. David Flockhart, a clinical pharmacologist, offered the least demanding standard for assessing a causal relationship. Dr. Flockhart acknowledged that epidemiologic and clinical trial data would be helpful, but found it to be unnecessary in assessing a causal relationship:

> Many important and well-recognized adverse drug reactions are not documented by well conducted epidemiologic studies that have been specifically conducted to detect them and that we still believe that the weight of the scientific evidence is sufficient to implicate their involvement to such an extent that we would remove a drug from the market. . . . Well conducted epidemiologic studies are of unquestionable value in this context when they are positive, but medical epidemiology is an inexact science, fraught with reliance on imperfect medical charts which are themselves really edited summaries of data available, or on ICD-9 codes or discharge diagnoses which may not reflect completely or accurately the actual course of events. As a result epidemiology is a somewhat blunt tool with which to ask detailed questions about adverse drug events and interactions, since the possible causes of type 2 errors: i.e. missing drug reactions that are actually there, are legion. Dr. Kulig used scientifically acceptable methodology in putting little reliance on the small amount of epidemiologic evidence available.[29]

Dr. Flockhart also noted that

> most clinical practice is not guided by data from [prospective, randomized, placebo-controlled clinical trials], because they are so difficult and expensive to conduct. In addition, aggregate data from such trials often do not apply to the specifics of individual cases. To assert that any medical practice has no scientific basis because a randomized, placebo-controlled trial to answer the pertinent question was not conducted would be to obviate the vast majority of clinical practice. It follows that other tools must usually be used to provide sufficient evidence to guide our practice.[30]

Instead, Dr. Flockhart reasoned by analogy that evidence of vasoconstriction in peripheral arteries caused by Parlodel and related substances would permit the inference that Parlodel would cause similar vasoconstriction in cerebral arteries, leading to increased

---

[28] *Id.* at 5-6.

[29] Letter from David Flockhart, 2-3 (Dec. 24, 2001) (on file in the law review office) [hereinafter Flockhart].

[30] *Id.* at 4-5.

blood pressure, intracranial hemorrhage, and stroke. Such reasoning presents an "analytical gap," in the words of *Joiner*, that far exceeds the modest stretch allowed by Dr. Savitz.

More specifically, Dr. Flockhart used a "weight of the evidence" approach to assess the effect of animal studies, case studies, and other available evidence.[31] Unlike Dr. Savitz, Dr. Flockhart relied on a series of case studies, including one case study reported in a peer-reviewed journal that recorded vasoconstriction changes in carotid arteries with the "challenge-dechallenge" (i.e., the introduction and removal) of a drug related to Parlodel (ergotamine tartrate):

> Case reports are often disparaged or discarded because they do not represent a plurality of events, but their strength is in making clear the *possibility* of an event, such as the possibility that Parlodel can cause vasoconstriction. In the determination of the possibility of drug involvement in an adverse reaction there is little that can substitute for a well informed investigative physician being present at the time of an event, performing a targeted series of diagnostic tests and then describing it carefully to us all via publication in the medical literature. A cause and effect relationship can be established even with a single case report if it is excellent. The demonstration of challenge and dechallenge is a particularly persuasive method in this context, and while this would be impossible in the case of stroke, it has been demonstrated for Parlodel for the closest reasonable surrogate: vasoconstriction.[32]

Dr. Flockhart also relied on a series of animal studies that indicated that Parlodel could cause vasoconstriction in peripheral arteries. After noting that doses are difficult to equate due to different metabolic rates, and that animals may have different pharmacologic receptors, Dr. Flockhart suggested that such studies

> can demonstrate the *possibility* of an effect, but they cannot carry the same weight as studies conducted in people or in human tissues. The data described in Exhibit 1013 indicating that Parlodel can cause superficial epithelial necrosis of dependant ear margins in dogs with long hanging ears demonstrate the *possibility* that the drug can bring about vasoconstriction in mammalian blood vessel in vivo even though the doses that were used were

---

[31] "To arrive at reasonable proof of a causal link in an individual case, one will commonly have to bring together *different* elements of evidence, none of which taken separately may be determinant, but which when viewed as a whole may be considered convincing." *Id.* at 2 (emphasis in original) (internal citation omitted); *see also* Vern R. Walker, *Risk Characterization and the Weight of the Evidence: Adapting Gatekeeping Concepts from the Courts*, 16 RISK ANALYSIS 793 (1996).

[32] Flockhart, *supra* note 30, at 3.

high.[33]

Dr. Flockhart then cited a number of published and unpublished references to the vascoconstrictive properties of Parlodel and other ergot derivatives and gave special weight to statements by the drug manufacturer that the risk of myocardial infarction or stroke exceeded the benefit of the drug for postpartum lactation, leading the FDA to withdraw its approval of Parlodel for this indication.

Dr. Flockhart acknowledged that none of the bits of evidence, considered individually, would be sufficient to support a finding of general causation, but "when viewed as a whole make a convincing case that Parlodel can cause stroke."[34]

Turning to the question of specific causation, Dr. Flockhart examined the differential diagnosis technique used by the plaintiff's two experts to rule out a number of plausible alternative causes of the stroke. Notably, Dr. Flockhart specifically did not exclude the possibility that the stroke was due to an elevated rate of unexplained stroke in postpartum women, since "it is unreasonable, and scientifically untenable, to hold that such rare strokes happen for no describable reason at all. It is likely that were each of these to be as closely examined as was Ms. Soldo's stroke, a scientifically plausible cause might well be found for each of them."[35]

While Dr. Flockhart found the differential diagnosis of Dr. Kulig to be scientifically acceptable, he found the differential diagnosis of Dr. Petro to fall short of the mark. Dr. Petro acknowledged that the cold medication Contac, which the plaintiff took, contains phenylpropanolamine (PPA) and may have been a contributing factor to the plaintiff's stroke, but offered no explanation for why he concluded that it was not the sole cause. According to Dr. Flockhart, this failure to exclude a specific plausible alternative cause through differential diagnosis rendered the methodology unreliable and inappropriate to the facts of the case.[36]

The third court-appointed expert, Dr. William J. Powers, a neurologist, offered an opinion that staked out an intermediate position regarding the requirements for recognizing "scientific knowledge." Dr. Powers was willing to tolerate an "analytical gap" somewhat larger than Dr. Savitz, but which fell short of that accepted

---

[33] *Id.* (emphasis in original).

[34] *Id.*

[35] *Id.* at 4.

[36] The expert report of Dr. Kulig does not mention this possibility, suggesting that he was not presented with this possibility.

by Dr. Flockhart. When Dr. Powers applied this standard to the proffered testimony of the plaintiff's experts, he too found it lacked sufficient scientific rigor.[37]

Dr. Powers also acknowledged that clinical or epidemiologic data are not necessary to declare a causal relationship, but objected to reasoning by analogy that evidence of vasoconstriction in peripheral and carotid arteries is indicative of a similar relationship in cranial arteries. In the absence of specific evidence, human or animal, that Parlodel causes vasoconstriction in cerebral arteries, he found the assertion of such an association to be without scientific foundation:

> Even more importantly, the review of evidence by Dr. Kulig does not consider the extensive scientific literature documenting that cerebral arteries respond differently to drugs of the ergot class than do peripheral vessels. Many studies indicate that ergots do not cause cerebral arterial vasoconstriction even though they do cause peripheral arterial vasoconstriction. The statement that cerebral vasoconstriction can cause intracerebral hemorrhage is not supported or substantiated by any observations or evidence in humans or animal models. As far as I am aware, no such data exists.[38]

Dr. Powers then applied the nine Bradford-Hill criteria for assessing causation to the proffered testimony, and found that it satisfied only one of the criteria—temporal sequence.[39]

Dr. Powers then assessed the differential diagnoses presented by the experts and found them lacking in scientific rigor. After noting that the differential diagnoses were appropriate and complete for ruling out a number of common causes of stroke, Dr. Powers noted that the process used to rule out other causes and implicate Parlodel "is logically flawed." He noted three flaws. First, the theory that intracerebral hemorrhage is part of a symptom of ergotism,[40] which

---

[37] Memorandum from William J. Powers, to Judge Donald J. Lee 1 (July 11, 2001) (on file with author) [hereinafter Powers].

[38] *Id.* at 2 (internal citation omitted).

[39] Here too there appears to be a difference in interpretation across the sciences, since Dr. Savitz indicated in his report that the Bradford-Hill criteria are to be used "as a means of interpreting an established association based on a body of epidemiologic research for the purpose of trying to judge whether the observed association reflects a causal relation between an exposure and a disease." Savitz, *supra* note 26, at 4.

[40] Ergotism is "a chronic poisoning produced by ingestion of ergot [a fungus that attacks plants], marked by cerebrospinal symptoms, spasm, cramps, a kind of dry gangrene of the extremities, and burning pain related to intense peripheral vasoconstriction." MILLER-KEANE ENCYCLOPEDIA & DICTIONARY OF MEDICINE, NURSING AND ALLIED HEALTH 471 (7th ed. 2000). Parlodel is an ergot derivative.

can be caused by Parlodel, is weakened by the fact that there were no symptoms of ergotism at the time of hospital admission for intracerebral hemorrhage. Second, the party's experts provided no adequate explanation for why Parlodel was more likely to cause the intracerebral hemorrhage than the Contac medication, which includes PPA, a product linked with intracerebral hemorrhage. Third, the party's experts provided no adequate reason for excluding the post-partum state itself as the cause.[41]

Having found that the testimony by the party's experts was not scientifically reliable, Dr. Powers then assessed the extent to which others in his field might express a contrary opinion that represents "a legitimate and responsible disagreement within [the] profession." Dr. Powers admitted that others outside his specialty of experimental pharmacology of the cerebral circulation might disagree because they are unfamiliar with the published scientific evidence regarding the difference in response of the cerebral vasculature and peripheral vasculature to ergot class drugs, and that such knowledge is typically not part of the training of specialists in the broader fields of neurology and cerebrovascular disease. Moreover, he noted that medical textbooks and peer-reviewed articles contain contrary views:

> The statement that cerebral vasoconstriction can cause intracerebral hemorrhage may be argued in the affirmative based on the occurrence of intracerebral hemorrhage in the conditions of hypertensive encephalopathy and declampsia and the belief that these conditions are caused by cerebral vasospasm. Although the best available evidence indicates that cerebral vasospasm is not present in hypertensive encephalopathy or eclampsia, this is still a commonly held belief that can be found in textbooks and review articles.[42]

In conclusion, Dr. Powers stated that in the absence of a clear cause for the hemorrhage, some measure of subjective judgment is required in assessing the evidence, and other persons, generally qualified in this field of expertise, may legitimately disagree:

> My conclusion that neither Dr. Kulig or Dr. Petro provides adequate justification to rule out other causes of intracerebral hemorrhage and implicate Parlodel as the primary cause is a subjective judgment based [on] my reading of the evidence and other persons, generally qualified in this field of expertise, may legitimately disagree because no other clear cause for the

---

[41] Powers, *supra* note 37, at 3.

[42] *Id.* at 6.

hemorrhage was established.[43]

B. The Court's Decision

One can imagine the disappointment, perhaps even regret, felt by the judge when he reviewed these court appointed experts' three independent and inconsistent decisions. It is likely that in calling for independent assessments from three different specialties, he hoped that the opinions would converge, thereby strengthening this assessment of the scientific methodology underlying the experts' opinions. Instead, he ended up with three opinions that shared little common ground.

Of course, the judge could have lessened the likelihood of disagreement by permitting the experts to consult together in reaching their opinions, as judges have in other cases.[44] Such consultation would have allowed each decision to be informed by the opinion of the others. The difference between Dr. Flockhart and Dr. Powers over the propriety of using research on noncranial arteries as a basis for inferring a similar effect on cranial arteries may have been resolved if they had the benefit of each other's thinking. Other conflicts, however, likely would have persisted, such as differences over the need for clinical or epidemiologic studies as a basis for a causal attribution in such cases. I believe the judge should be commended for soliciting independent views without consultation, since these views give a more accurate assessment of the variation of opinion that exists within the various branches of science. Judges should resolve such conflicts within the constraints of legal doctrines, rather than encourage the experts to negotiate among themselves to find a common resolution that they can present to the court. In fact, it is the resolution of the conflicts among the legitimate views across the sciences that, in my opinion, poses the greatest difficulty to the faithful application of the gatekeeping responsibility for expert testimony.

At first glance, one might assume that three different opinions from highly qualified experts appointed by the court, in itself, would be evidence of "a legitimate and responsible disagreement" within the relevant sciences regarding material issues and requiring that the disputed issues be presented to a jury. Moreover, the two court-appointed experts who found the plaintiff's testimony lacking in scientific integrity acknowledged the existence of "legitimate and

---

[43] *Id.*

[44] Hooper et al., *supra* note 20, at 161-64.

responsible disagreement" within the profession.[45]

Sanders presumably would allow the jury to consider such different views, assuming the jury is capable of comprehending the differences in the views between the parties' experts and appointed experts and resolving these differences in a thoughtful and reasoned way. If the purpose of restrictions on admissibility is to "shelter jurors from their own shortcomings" when they are unable to distinguish between good and bad testimony and will, therefore, "be more likely to reach an incorrect conclusion,"[46] is there a need for such protection when the breadth of legitimate scientific views is as great as indicated by the opinions of the court-appointed experts? It is hard to imagine that the jury could not find a defensible decision within this span of views.

The court, however, took no comfort in the diversity of views presented by the court-appointed experts. Faced with such conflicting opinions, the court attempted to reconcile disputes over scientific validity that the scientific community itself had shied away from. In doing so, the court unavoidably stepped beyond the bounds of assessing whether the methodology and techniques employ the same degree of intellectual rigor appropriate to the profession and instead established legal policy regarding how such disputes among scientists are to be reconciled under law.

In an opinion that extends more than 100 pages, the court dissected the views of Dr. Flockhart and the plaintiff's experts, and concluded that the opinions expressed by the plaintiff's experts must be excluded since they "failed to use a reliable scientific methodology" to demonstrate general causation and specific causation. Having excluded the testimony of the plaintiff's experts, the court then granted summary judgment in favor of the defendant.[47]

---

[45] Dr. Savitz did not answer precisely the question regarding the extent to which a contrary opinion "would likely be viewed by others in the field as representing legitimate and reasonable disagreement within your profession." Instead, he indicated that "80% of my peers would concur" with his determination, leaving 20% who may disagree. Savitz, *supra* note 25, at 5. Dr. Savitz also notes that "most scientists would likely share the view that leaping across the huge gulf of critical data moved the person making the inference beyond the scope of scientific assessment." *Id.* at 6. Dr. Powers notes that a contrary view "is still a commonly held belief that can be found in textbooks and review articles," suggesting that the contrary view may even be generally accepted. Powers, *supra* note 37, at 6.

[46] Sanders, *supra* note 1, at 891.

[47] In doing so, the court "adopted almost verbatim most of the proposed findings of fact and conclusions of law submitted by the defendant for the reason those proposed findings of fact and conclusions of law correctly reflect the facts in the

In considering the testimony of the plaintiff's experts, the court indicated "differential diagnosis alone cannot establish causation to a degree of medical certainty in a case involving a disease as common as stroke."[48] Given that approximately one-third of all strokes go undiagnosed as to their cause, "the scientific way to determine whether bromocriptine increases the risk of stroke in humans is through a proper controlled clinical or epidemiologic study."[49]

The court's examination of Dr. Flockhart views is particularly informative of the manner in which many federal courts consider expert testimony. It examined each element of evidence on which Dr. Flockhart based his opinion, and pointed out that each one demonstrates only the *possibility* that Parlodel can cause intracranial hemorrhage and stroke. The court made clear that epidemiology studies or clinical trials are not essential to establishing causation, but endorsed Dr. Savitz's view that the absence of such studies would require "a tremendous amount of indirect evidence" to establish causation.[50] As an example, the court mentioned that "a scientifically valid understanding of the alleged mechanism by which Parlodel allegedly causes vasoconstriction could be important indirect evidence, but it [did] not exist in this case."[51] At a later point, the court noted that such a linkage must be obvious, "in the way that the tornado leading to injury is obvious."[52]

The court then dismissed the clinical case reports, since they demonstrated only the possibility of a causal relationship. As to the published case report that demonstrates vasoconstriction in the presence of a drug closely related to Parlodel using a challenge/dechallenge methodology, the court noted that the authors suggested only the "possible" causal relationship, and that the case report demonstrated only coronary vasoconstriction and not vasoconstriction of cerebral arteries or stroke.[53]

The court then dismissed the animal studies that Dr. Flockhart and the plaintiff's experts relied on, noting again that such studies demonstrated only the possibility of a causal relationship and do not explain adequately how such data relate to humans. The court was

---

record as well as relevant law." *Soldo*, 2003 WL 355931, at *2 n.2.

[48] *Id.* at *9.

[49] *Id.*

[50] *Id.* at *72.

[51] *Id.*

[52] *Id.* at *74.

[53] *Soldo*, 2003 WL 355931, at *83. Of course, there can be no challenge/dechallenge demonstration of a stroke, since the effect of a stroke is not reversible.

especially skeptical of a study of vasoconstriction in the peripheral arteries of the ears of dogs:

> [D]espite the fact that Dr. Flockhart and plaintiff's experts recognize that human studies carry greater weight than animal studies, they provide no explanation for why they give more weight to an animal study showing alleged effects in the "dependent ear margins in dogs with long hanging ears" than negative human studies or human studies demonstrating vasodilation, given that plaintiff is not a dog and does not have long hanging ears."[54]

The court then turned to the distinction raised by Dr. Powers between the reaction of cerebral arteries and other vascular systems, and noted:

> Plaintiff experts who claim that Parlodel acts like all other ergots to allegedly cause vasoconstriction, provide no reliable means of distinguishing studies showing that cerebral arteries respond differently to ergots than do peripheral vessels [citation omitted] . . . . Studies showing that cerebral arteries respond differently to ergots than do peripheral vessels invalidate plaintiff's expert's efforts to reason by analogy from scientific data regarding alleged vasoconstriction of the peripheral vessels.

The court then criticized the experts for failing to rule out a number of other possible causes, such as caffeine, smoking, stress, hormones, PPA exposure due to Contac cold medication, and blood abnormalities. After observing that the plaintiff's experts failed to rule out the postpartum period or idiopathic stroke as possible causes of the intracranial hemorrhage, the court then questioned Dr. Flockhart's approval of this practice. Noting that Dr. Flockhart asserted that "[i]t is likely that were each of these [postpartum intracranial hemorrhages] to be as closely examined as was Ms. Soldo's[,] that a scientifically plausible cause [other than the postpartum period] might well be found for each of them."[55] The court then stated that Dr. Flockhart offered no support for these statements other than his own *ipse dixit*.[56]

The court did not explicitly address Dr. Flockhart's "totality of the evidence" test, but did not appear to give weight to the fact that different kinds of evidence point in the same direction. Quoting Dr. Savitz, the court noted that while expert opinions may make

> "appropriate use of all of the available information, [] in the

---

[54] *Id.* at *77 (internal citations omitted).
[55] *Id.* at *87.
[56] *Id.*

> absence of some minimum amount or level of scientific evidence, the opinions cannot be scientifically derived because there is too little science from which to derive them."
>
> Although it is sometimes necessary in a clinical, regulatory, or business practice to make decisions based on less than sufficient and/or reliable scientific evidence due to practical demands requiring immediate decision-making, such guesses, although perhaps reasonable hypotheses based on the best available evidence, do not constitute a scientifically reliable approach when used to assess causality via the scientific method.[57]

In conclusion, the court voiced its agreement with Drs. Powers and Savitz:

> The body of scientific evidence relating to Parlodel and stroke is simply insufficient to support a scientifically reliable application of plaintiff's expert methodology. . . . Without sufficient reliable evidence of general causation, plaintiff's experts could not reliably apply a differential diagnosis that comports with the scientific method, notwithstanding the fact that physicians in clinical practice may be required to proceed with a differential diagnosis on the basis of guesses or hypotheses due to the exigency of the need to treat their patients.[58]

II. THE QUEST FOR *IPSE DIXIT*

Soldo reveals what other decisions have obscured: in making a gatekeeping determination the court must reconcile conflicting values of numerous sciences, each with differing intellectual processes, differing assumptions, and differing degrees of tolerance for extrapolation from scientific studies to human circumstances. Conflicts among the three court-appointed experts could not be attributed to party sponsorship or adversarial presentation. Even when free of distortions imposed by the legal forum, distinguished scholars from different disciplines will invoke diverse standards and practices in assessing evidence and may reach divergent conclusions regarding the presence or absence of a causal relationship.

In resolving such conflicts through an admissibility determination, courts do more than exert paternalist vigilance to compensate for juries' perceived shortcomings. It is not clear, in light of the research cited by Sanders, that a jury would be incapable of understanding the different positions of the parties' experts and reaching a reasonable decision. By dismissing the report by the

---

[57] *Id.* at *69-70 (internal citations omitted).

[58] *Id.* at *95.

court-appointed clinical pharmacologist as failing to meet a sufficient standard of scientific reliability, the court was in fact establishing a legal threshold for sufficiency that was independent of any uniform scientific standard. The court made clear that the fact that a clinical pharmacologist employs "in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field" is not sufficient for consideration by a jury. For the *Soldo* court and a good many others, admissible evidence requires more than meeting appropriate standards of professional practice; it also requires demonstration of a relationship through methodologies that are not an essential part of clinical practice.

The court in *Soldo* indicated as much when it noted that even if the plaintiff's experts' opinions were admissible under *Daubert*, "such evidence provides but a scintilla of support for plaintiff's position and would not be sufficient to allow a reasonable jury to find that plaintiff's ICH had been caused by Parlodel."[59] One might argue that this is a more appropriate basis for such a decision, rather than striking the testimony as inadmissible due to some perceived flaw in the methodology and reasoning. Even if the experts use methods and reasoning appropriate to their profession, the courts may set, as a matter of law, a minimum threshold for evidence that is sufficient to justify submission to a jury. Of course, if such a standard is established in an explicit manner as insufficient as a matter of law, that decision would be subject to appellate review on a *de novo* basis and therefore more vulnerable to reversal on appeal. Such an approach, however, would seem preferable to declaring that broadly approved professional practices are scientifically unsound, and therefore, do not meet the standards for admissibility.

Coming to terms with the role of the courts in setting such sufficiency standards will not be easy. As Cohen points out, in these circumstances the vocabulary of science may obscure as much as it reveals.[60] Courts utilizing scientific terms assign them different meanings, thereby obscuring the values that are being endorsed and diverting scientists who seek to appear as experts into meaningless debates about whether their declarations are "scientifically valid." The courts cannot resolve the diverse views and values that characterize the scientific academy; one science's accepted methodology may be another science's *ipse dixit*. A court can, however, specify a minimum threshold for admissible scientific evidence, and make clear that in doing so it is establishing a legal

---

[59] *Id.* at *93.

[60] Cohen, *supra* note 2, at 944-45; *see also id.* Parts II & III.

standard and not assessing the ephemeral concept of "scientific validity.